

# Allelic Distribution at Microsatellite Xuhw89 locus, a Functional SSR Marker for Grain Protein Content in Pakistani Wheat Germplasm

A. KHALID<sup>1,2\*</sup> and A. HAMEED<sup>1,2</sup>

<sup>1</sup>Nuclear Institute for Agriculture and Biology (NIAB), Jhang Road, Faisalabad, Pakistan

<sup>2</sup>Pakistan Institute of Engineering and Applied Sciences, Islamabad, Pakistan

(Received 18 February 2019; Accepted 21 May 2019;  
Communicated by F. Békés)

Grain Protein Content (GPC) of wheat is significant for enhanced nutritional value and is one of the most important aspects effecting pasta and bread making quality as well. Seventy seven Pakistani wheat varieties and advance lines were analysed to access the allelic distribution at microsatellite Xuhw89 locus using functional SSR marker. Overall, 42% of tested wheat genotypes were found to carry 126+130-bp allele while a 126-bp allele was detected in 58% of genotypes. A target band of 126-bp was amplified in all tested genotypes, however, an additional band of 130-bp was also detected along with 126-bp band in 32 genotypes. Genotypes i.e. (Punjab-96) total soluble protein and (MEXI PAK) globulin with 126+130-bp allele while (Faisalabad-2008) salt soluble protein and (TC-4928) albumin with allele of 126-bp depicted highest grain protein content. The alleles identification associated with maximum grain protein content in Pakistani wheat germplasm will assist in accelerating the breeding program in future.

**Keywords:** wheat, grain protein content, SSR marker, Xuhw89

## Introduction

Bread wheat (*Triticum aestivum*) is the principal cereal crop worldwide with annual production of about 700 million tons thus fulfilling 20% of overall calorie consumption throughout world (Vishwakarma et al. 2016; Vishwakarma et al. 2014). Protein is an essential nutrient for the survival of humans and animals. Proteins found especially in mature cereal grains are responsible for providing considerable amount of world's plant protein, and its proportion defines the nutritional quality in addition to end use characteristics of the grain (Fan et al. 2017). It is fundamental constituent of cells as it supports immune as well as enzymatic responses besides muscle development and its malnutrition results in marasmus or kwashiorkor (Krishnappa et al. 2017).

Grain protein content (GPC) partly governs the nutritional significance as well as the baking attributes of *Triticum aestivum*. A composite genetic system regulates this typical quantitative character and is influenced by environmental aspects as well as management practices, i.e. light intensity, nitrogen availability, water accessibility, and temperature.

\*Corresponding author; E-mail: [invincible\\_me2@yahoo.com](mailto:invincible_me2@yahoo.com)

(Blanco et al. 2012). Elevated protein level has also customarily been associated with a reduced yield potential. (Turner et al. 2004).

The Gpc-B1 allele for high grain protein content was initially recognized in wild emmer wheat *Triticum turgidum* ssp. *dicoccoides* accession FA15-3 known as DIC (Joppa and Cantrell 1990) developed of the DIC chromosomes substitution lines development in the cultivar 'Langdon' (LDN) disclosed a locus for high grain protein content on chromosome 6B (Uauy et al. 2006). For the wheat varieties to be utilized in bread making, the high grain protein level greater than 12.5% is essential for augmenting export potential of wheat and for improved diet (Turner et al. 2004; Vishwakarma et al. 2014).

Grain protein content improvement and change in bread wheat storage proteins composition have been the main concern for plant breeders. The choice for enhanced kernel protein content or for elevated amount of few individual necessary amino acids/protein subunits though, is costly and laborious. Evolution of such molecular markers which enables indirect marker assisted selection (MAS) for grain protein content is a suitable substitute. This indirect (MAS) is of primary importance for characters such as kernel protein content, which show increased genotype  $\times$  environment interaction resulting in decreased heritability. In current studies, some molecular markers linked with grain protein content in tetraploid and hexaploid wheat have been recognized (Prasad et al. 2003).

Classification of genetic variation and relatedness is a basic component in the improvement of crop strategies (Mohammadi and Prasanna 2003). A lot of approaches are presently employed for genetic diversity analysis in germplasm breeding lines, accessions and segregating populations. These techniques were centred on parentage, morphological as well as agronomic performance and biochemical/molecular data. Morphological features used for the evaluation of genetic variation are frequently influenced by environmental conditions, whereas the molecular markers expression being the direct genes product remains uninfluenced. Moreover, by using molecular markers, various germplasm lines might be categorized in a small time period and reveal more precise genetic diversity (Mahmood et al. 2011). Therefore, molecular markers provide comprehensive genetic resources characterization and offer a direct measure of genetic variability besides, go afar the indirect diversity measures based on geographic origin or agronomic attributes.

Microsatellites are simple sequence repeats (SSR) of 1 to 6 nucleotides. They are ample, distributed all over the genome and depict greater polymorphism level as compared to other genetic markers. These aspects together with their ease of detection have proved them valuable markers. Their automation potential and inheritance in a co-dominant mode are further benefits in comparison with other markers. SSR markers cover entire twenty-one wheat chromosomes. They have been employed to classify genetic variability in wild relatives as well as in a seed bank pool of upgraded wheat germplasm (Drikvand et al. 2015; Islam et al. 2012)

Hence, the objective of current study is to analyse and catalogue allelic variants at Xuhw89 locus tightly linked to the grain protein content gene *Gpc-B1* using simple sequence repeat (SSR) technique in Pakistani wheat genetic resource.

## Materials and Methods

### *Plant material collection*

Wheat germplasm of 77 varieties were collected from different centres in each agro climatic region of Pakistan (Khalid and Hameed 2017) for the analysis of allelic variation at Xuhw89 locus tightly linked to the grain protein content gene *Gpc-B1*.

### *DNA extraction, PCR amplification and molecular marker analysis*

Total genomic DNA was extracted from ten days old seedlings by the method of (Plaschke et al. 1995). The quality and quantity of DNA was determined by spectrophotometer and agarose gel electrophoresis. The primer pair Xuhw89 used for PCR amplification was reported by (Distelfeld et al. 2005). PCR reactions were performed according to (Roder et al. 1998) in 25 µl reaction volume encompassing the following constituents: 100 ng of template DNA, 0.2 mM of each of the four dNTPs, 1x Taq polymerase buffer, 1 unit Taq polymerase, 1.5 mM MgCl<sub>2</sub> and 0.4 µM each of the two primers. Amplifications were performed in a BioRad gradient thermal cycler. The amplification cycles consisted of 40 cycles with denaturation at 94 °C for 2 minute, annealing at 58 °C for 30 sec and extension at 72 °C for 1 min. The final extension was done at 72 °C for 10 min.

Six per cent non-denaturing polyacrylamide gels were used to separate amplification products and fragments were identified by ethidium bromide staining. Gels were documented using UVI pro Platinum 1.1 System and UVI Band Map software (UVItec Ltd., Cambridge, UK). The generated data matrix was subjected to statistical analyses using XL STAT software and was analysed by means of two-way ANOVA with replications. Significance of data was tested by analysis of variance and Tukey (HSD) Test at  $p < 0.05$  and where applicable at  $p < 0.01$ .

## Results

Molecular analysis of seventy-seven Pakistani wheat genotypes revealed that Xuhw89 was polymorphic marker and detected allelic variants of *Gpc-B1* gene in Pakistani wheat germplasm. All the samples successfully amplified the marker Xuhw89. The electrophoretic patterns of microsatellite locus Xuhw89 depicted that target band of 126-bp was amplified in all tested genotypes, however, an additional band of 130-bp was also detected along with 126-bp band in 32 genotypes (Figs S1\*, S2, S3 and S4). In total, two different alleles, i.e. 126-bp and 130-bp were detected at the Xuhw89 locus in tested wheat genotypes (Table 1).

After the identification of alleles, estimates of percentage distribution of protein content in different wheat genotypes specific alleles for the microsatellite locus Xuhw89 were calculated. Overall, 42% of tested wheat genotypes were found to carry 126 + 130-bp allele (Fig. 1) while a 126-bp allele was detected in 58% of genotypes.

\*Further details about the Electronic Supplementary Material (ESM) can be found at the end of the article.

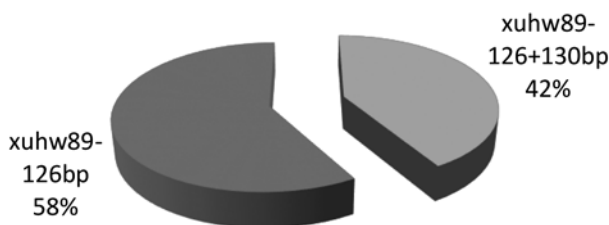


Figure 1. Percentage distribution of detected alleles for the microsatellite locus Xuhw89 in Pakistani wheat genotypes

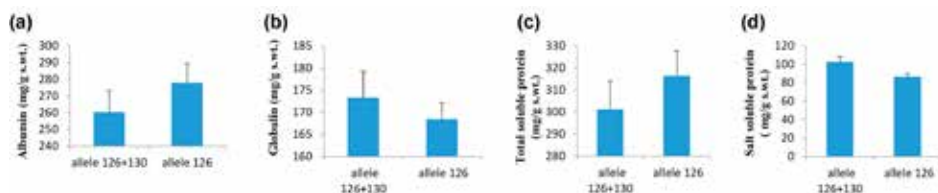


Figure 2. Mean values for albumin (a), globulin (b), total soluble protein (c) and salt soluble protein (d) in different wheat genotypes

Mean values for albumin (Fig. 2a), globulin (Fig. 2b), total soluble protein (Fig. 2c) and salt soluble protein (Fig. 2d) for genotypes having 126-bp allele and 126 + 130-bp allele were calculated and compared. Results depicted that albumin was higher in genotypes with 126 bp allele ( $277.921 \pm 4.061$  mg/g s. wt.) while globulin was found maximum in genotypes with 126 + 130 bp allele ( $173.313 \pm 5.997$  mg/g s. wt.). Maximum total soluble protein as well as salt soluble protein was detected in genotypes with 126 bp allele ( $316.400 \pm 11.349$  mg/g s. wt.) and 126 + 130 bp allele ( $102.778 \pm 5.716$  mg/g s. wt.) respectively.

Table 1. Name of genotypes and allele size at microsatellite locus Xuhw89 in Pakistani wheat varieties

Sr. No.	Genotypes	Alleles (bp)
1.	PERWAZ, PAVON, LU-26, AUQAB-2000, Bhakkar-2000, AS-2002, SH-2002, Manthar-2003, Shafaq-2006, Fareed-2006, Chakwal-50, BARS-2009, Sehar-2006, Dharabi 2011, Faisalabad-2008, AARI-2011, Millat-2011, Punjab-2011, Tatara 1996, Sindh-81, Soghat-90, Kiran-95, 2005, Nia Amber, Punjab-90, Nifa Lalma 2012, Jauhar-78, Barani-83, Sassui-2006, C-591, 2006, MEXI PAK	126 + 130
2.	Margalla-99, Mehran-89, Pari-73, Saleem-2000, SA-75, Nesser, S-24, NR-421, Zardana, Raskoh-2005, Pasban-90, Sitta, WL-711, Nifa-Bathoor, Punjab-96, 2156, Kohistan, 9021, TC-4928, 6544-6, 9244, Suleman, NR-234, NARC-2011, Watan-94, Marvi, Abadghar-93, C-228, Benazir-12, Nia Sunhari, Nia Sundar, IQBAL-2000, UFAQ 2002, Lasani-2008, Mairaj-2008, Inqulab-91, Galaxy-2013, Bakhtawar-1993, Fakhar-e-Sarhad, Ujala-16, Takbeer-2000, Sarsabz, Bhittai-2004, GA-2002, Khirman 2006	126

Table 2. Descriptive statistics for differential wheat grain protein contents in genotypes with different Xuhw89 alleles

Differential proteins/ Xuhw89 locus	Min.	Max.	Range	Mean	SD	Variation coefficient	SE
Albumin A126+130	199.556	326.889	127.333	260.333	31.007	0.117	5.481
Albumin   A-126	222.222	<b>352.889</b>	130.667	277.921	27.243	0.097	4.061
Globulin A126+130	112.444	<b>252.667</b>	140.222	173.313	33.922	0.193	5.997
Globulin  A-126	126.667	226.667	100.000	168.504	24.010	0.141	3.579
TSP  A-126+130	<b>68.667</b>	462.667	394.000	301.278	72.480	0.237	12.813
TSP   A-126	207.778	487.333	279.556	316.400	76.134	0.238	11.349
SSP   A-126+130	61.333	162.444	101.111	102.778	32.333	0.310	5.716
SSP   A-126	63.778	161.111	97.333	86.499	23.514	0.269	3.505

SD = standard deviation, SE = standard error, TSP = total soluble protein, SSP = salt soluble protein, A = allele.

In genotypes with 126 + 130 bp allele albumin ranged (127.33) from a minimum value of 199.56 to a maximum value of 326.89 while in genotypes with 126 bp allele albumin ranged (130.67) from a minimum value of 222.22 to a maximum value of 352.89 (Table 2). In genotypes with 126 + 130-bp allele globulin ranged (140.22) from a minimum value of 112.44 to a maximum value of 252.67 whereas in genotypes with 126-bp allele globulin ranged (100.0) from a minimum value of 126.67 to a maximum value of 226.67. Albumin and globulin revealed maximum difference in genotypes with 126-bp allele and 126 + 130-bp allele, respectively. In genotypes with 126 + 130-bp allele total soluble protein ranged (394.0) from a minimum value of 68.67 to a maximum value of 462.67 however, in genotypes with 126-bp allele total soluble protein ranged (279.56) from a minimum value of 207.78 to a maximum value of 487.33. In genotypes with 126 + 130-bp allele salt soluble protein ranged (101.11) from a minimum value of 61.33 to a maximum value of 162.44 while in genotypes with 126-bp allele globulin ranged (97.33) from a minimum value of 63.78 to a maximum value of 161.11. Total soluble protein depicted maximum difference in genotypes with allele 126 + 130-bp.

## Discussion

Wheat has significant importance with respect to production and consumption among all cereals. Protein proportion in grain of bread wheat play a pivotal role in defining end use quality and holds a substantial role in human diet as high protein has been stated to increase nutritive along with bread-making quality (Mallick et al. 2013). Recently, the wheat quality has been a topic of discussion, as greater yield has made a trend towards decreased concentration of protein and this makes it necessary to identify alleles using molecular markers that affect grain protein content. To our knowledge, identification of allelic variants of Xuhw89 locus for *GPC-B1* gene in Pakistani wheat genetic resource

has not been executed and for the first time genetic diversity of genotypes regarding above-mentioned locus/gene have been performed.

Albumin and globulin makes 20 to 25% of total protein in wheat endosperm. Both are water miscible, biologically active, holds allergenic effects and play essential part in metabolism as well as nutrition (Khan et al. 2013; Khan et al. 2015; Merlino et al. 2009).

In general, albumin was found higher in genotypes with 126-bp allele ( $260.333 \pm 5.481$  mg/g s. wt.), but specifically found maximum in advance line TC-4928 ( $352.89$  mg/g s. wt.) with 126-bp allele, during current study. Maximum globulin was found in genotypes with 126 + 130-bp allele ( $173.313 \pm 5.997$  mg/g s. wt.) on the whole, while it was found highest specifically in genotype MEXI PAK ( $252.67$  mg/g s. wt.) with 126 + 130-bp allele.

Wheat seed storage proteins are responsible for estimating quality as flexibility and extensibility of dough is due to them (Rasheed et al. 2014). Over all, maximum total soluble protein was detected in genotypes with 126-bp allele ( $316.400 \pm 11.349$ ) however, precisely, it was observed maximum in Faisalabad-2008 ( $162.44$  mg/g s. wt.) with 126-bp allele. Highest salt soluble protein was identified in genotypes with 126 + 130-bp allele ( $102.778 \pm 5.716$ ) in general, but specifically it was observed maximum in Punjab-96 ( $487.33$  mg/g s. wt.) with 126 + 130-bp allele. Generally, 42% of tested wheat genotypes were found to carry 126 + 130-bp allele while a 126-bp allele was detected in 58% of genotypes.

Current study revealed that selected wheat varieties from various geographical zones of Pakistan, viz. (Punjab-96) total soluble protein and (MEXI PAK) globulins with allele of 126 + 130 bp while (Faisalabad-2008) (salt soluble protein) and (TC-4928) albumins with 126 bp allele have diversity with respect to *GPC-B1* locus. Occurrence and frequency of *GPC-B1* gene at DNA level will aid in precise parents selection intended for wheat breeding program.

### Acknowledgements

Authors are thankful for provision of facilities and help in analysis at the Nuclear Institute for Agriculture and Biology MAB-Lab1, Faisalabad. The paper is part of PhD thesis of first author i.e. Anam Khalid.

### References

- Blanco, A., Mangini, G., Giancaspro, A., Giove, S., Colasuonno, P., Simeone, R., Signorile, A., De Vita, P., Mastrangelo, A., Cattivelli, L. 2012. Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. *Mol. Breeding* **30**(1):79–92.
- Distelfeld, A., Uauy, C., Fahima, T., Dubcovsky, J. 2005. Physical map of the wheat high grain protein content gene *Gpc-B1* and development of a high throughput molecular marker. *New Phytol.* **169**(4):753–763.
- Drikvand, R., Najafian, G.E., Salahvarzi, A. 2015. Investigation of genetic diversity of some durum and bread wheat genotypes using SSR markers. *J. Biodivers. Environ. Sci.* **6**(3):24–32.

- Fan, C., Zhai, H., Wang, H., Yue, Y., Zhang, M., Li, J., Wen, S., Guo, G., Zeng, Y., Ni, Z. 2017. Identification of QTLs controlling grain protein concentration using a high-density SNP and SSR linkage map in barley (*Hordeum vulgare* L.). *BMC plant Biol.* **17**(1):122.
- Islam, S., Haque, M., Emon, R., Islam, M., Begum, S. 2012. Molecular characterization of wheat (*Triticum aestivum* L.) genotypes through SSR markers. *Bangladesh J. Agr. Res.* **37**(3):389–398.
- Joppa, L., Cantrell, R. 1990. Chromosomal location of genes for grain protein content of wild tetraploid wheat *Crop Sci.* **30**(5):1059–1064.
- Khalid, A., Hameed, A. 2017. Seed biochemical analysis based profiling of diverse wheat genetic resource from Pakistan. *Front. Plant. Sci.* **8**:1–15.
- Khan, S., Ghanghro, A.B., Memon, A., Tahir, I., Shah, A.M., Sahito, M.A., Talpur, F.N., Qureshi, S. 2013. Quantitative analysis of wheat proteins in different varieties grown in Sindh, Pakistan. *Intl. J. Agri. Crop Sci.* **5**(16):1836–1839.
- Khan, S., Memon, A., Ghanghro, A., Nabi, G. 2015. Characterization of wheat protein (Albumin) in different varieties of wheat cultivated in Sindh through SDS-PAGE electrophoresis. *Sindh Univ. Res. J.* **47**(2):361–366.
- Krishnappa, G., Singh, A.M., Chaudhary, S., Ahlawat, A.K., Singh, S.K., Shukla, R.B., Jaiswal, J.P., Singh, G.P., Solanki, I.S. 2017. Molecular mapping of the grain iron and zinc concentration, protein content and thousand kernel weight in wheat (*Triticum aestivum* L.). *PloS One* **12**(4):e0174972.
- Mahmood, T., Siddiqua, A., Rasheed, A., Nazar, N. 2011. Evaluation of genetic diversity in different Pakistani wheat land races. *Pak. J. Bot.* **43**(2):1233–1239.
- Mallick, S., Azaz, K., Gupta, M., Sharma, V., Sinha, B. 2013. Characterization of grain nutritional quality in wheat. *Indian J. Plant. Physi.* **18**(2):183–186.
- Merlino, M., Leroy, P., Chambon, C., Branlard, G. 2009. Mapping and proteomic analysis of albumin and globulin proteins in hexaploid wheat kernels (*Triticum aestivum* L.). *Theor. Appl. Genet.* **118**(7):1321–1337.
- Mohammadi, S., Prasanna, B. 2003. Analysis of genetic diversity in crop plants – salient statistical tools and considerations. *Crop Sci.* **43**(4):1235–1248.
- Plaschke, J., Ganal, M., Röder, M. 1995. Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theor. Appl. Genet.* **91**(6):1001–1007.
- Prasad, M., Kumar, N., Kulwal, P., Röder, M., Balyan, H., Dhaliwal, H., Gupta, P. 2003. QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. *Theor. Appl. Genet.* **106**(4):659–667.
- Rasheed, A., Xia, X., Yan, Y., Appels, R., Mahmood, T., He, Z. 2014. Wheat seed storage proteins: Advances in molecular genetics, diversity and breeding applications. *J. Cereal. Sci.* **60**(1):11–24.
- Röder, M.S., Korzun, V., Gill, B.S., Ganal, M.W. 1998. The physical mapping of microsatellite markers in wheat. *Genome* **41**(2):278–283.
- Turner, A., Bradburne, R., Fish, L., Snape, J. 2004. New quantitative trait loci influencing grain texture and protein content in bread wheat. *J. Cereal. Sci.* **40**(1):51–60.
- Uauy, C., Brevis, J.C., Dubcovsky, J. 2006. The high grain protein content gene *Gpc-B1* accelerates senescence and has pleiotropic effects on protein content in wheat. *J. Exp. Bot.* **57**(11):2785–2794.
- Vishwakarma, M.K., Arun, B., Mishra, V., Yadav, P., Kumar, H., Joshi, A.K. 2016. Marker-assisted improvement of grain protein content and grain weight in Indian bread wheat. *Euphytica* **208**(2):313–321.
- Vishwakarma, M.K., Mishra, V., Gupta, P., Yadav, P., Kumar, H., Joshi, A.K. 2014. Introgression of the high grain protein gene *Gpc-B1* in an elite wheat variety of Indo-Gangetic Plains through marker assisted back-cross breeding. *Curr. Plant Biol.* **1**:60–67.

### Electronic Supplementary Material (ESM)

Electronic Supplementary Material (ESM) associated with this article can be found at the website of CRC at <https://akademai.com/loi/0806>

Electronic Supplementary *Figure S1*. Xuhw89 maker amplification showing the presence of different allelic variants in Pakistani wheat germplasm. M = 50 bp DNA ladder; lane 1: Pavon; lane 2: PARWAZ; lane 3: MEXI-PAK; lane 4: LU-26; lane 5: IQBAL 2000; lane 6: AUQAB 2000; lane 7: UFAQ 2002; lane 8: GA 2002; lane 9: Bhakkar 2000; lane 10: SH-2002; lane 11: AS-2002; lane 12: Manthar 2003; lane 13: Fareed 2006; lane 14: Shafaq 2006; lane 15: Mairaj 2000; lane 16: Lasani 2008; lane 17: Inqulab 91; lane 18: Sehar 2006; lane 19: chakwal 50; lane 20: BARS-2009; M = 50 bp DNA ladder

Electronic Supplementary *Figure S2*. Xuhw89 maker amplification showing the presence of different allelic variants in Pakistani wheat germplasm. M = 50 bp DNA ladder, lane 21: Dharabi 2011; lane 22: Faisalabad 2008; lane 23: AARI 2011; lane 24: Millat 2011; lane 25: Punjab 2011; lane 26: Galaxy 2013; lane 27: V-8203; lane 28: Bhaktawar 1993; lane 29: Tatar; lane 30: Fakhr e Sarhad; lane 31: Takbeer 2000; lane 32: nia lalma 2013; lane 33: Jauhar 78; lane 34: Sindh-81; lane 35: Sarsabz; lane 36: Soghat 90; lane 37: Kiran 95; lane 38: Bhattai 2004; lane 39: Barani 83; lane 40: sassui, M = 50 bp DNA ladder

Electronic Supplementary *Figure S3*. Xuhw89 maker amplification showing the presence of different allelic variants in Pakistani wheat germplasm. M = 50 bp DNA ladder, lane 41: Khirman 2006; lane 42: nia amber; lane 43: nia sunhari; lane 44: nia sundar; lane 45: Benazir-12; lane 46: NARC 400; lane 47: Watan-94; lane 48: Marvi; lane 49: Abadghar; lane 50: C-228; lane 51: C-591; lane 52: 2006; lane 53: 2005; lane 54: 2156; lane 55: Kohistan; lane 56: 9021; lane 57: TC-4928; lane 58: 6544-6; lane 59: 9244; lane 60: Suleman; M = 50 bp DNA ladder

Electronic Supplementary *Figure S4*. Xuhw89 maker amplification showing the presence of different allelic variants in Pakistani wheat germplasm. M = 50 bp DNA ladder, lane 61: NR-234; lane 62: NR-421; lane 63: Margalla-99; lane 64: sitta; lane 65: Nesser; lane 66: Zardana; lane 67: Raskoh 2005; lane 68: Mehran-81; lane 69: Saleem 2000; lane 70: Punjab 90; lane 71: SA-75; lane 72: Pasban 90; lane 73: Nifa-Bathoor; lane 74: Punjab 96; lane 75: S-24; lane 76: Pari-73; lane 77: WL-711